

## Introspection, déploiement et simulation

Jérôme Dokic

Volume 32, numéro 2, automne 2005

URI : <https://id.erudit.org/iderudit/011874ar>

DOI : <https://doi.org/10.7202/011874ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer cet article

Dokic, J. (2005). Introspection, déploiement et simulation. *Philosophiques*, 32(2), 383–397. <https://doi.org/10.7202/011874ar>

### Résumé de l'article

Selon une théorie cognitiviste de l'auto-attribution, je peux parvenir à la connaissance directe, non-inférentielle de mes propres croyances. Cette théorie a été traditionnellement associée à la notion d'introspection conçue comme source de connaissance interne. On sait (au moins depuis Wittgenstein) que le recours à cette notion compromet l'application à soi-même d'un concept unifié de croyance, valable également pour autrui. Dans cet essai, j'explore une autre méthode d'auto-attribution, également envisagée par Wittgenstein (et plus tard par Gareth Evans), que j'appelle « méthode de déploiement ». Selon cette méthode, je parviens à la connaissance de mes croyances en portant mon attention, non pas à l'intérieur de moi-même, mais directement sur le monde extérieur tel que je l'ai trouvé. Certains arguments wittgensteiniens suggèrent que la méthode de déploiement conduit inexorablement au solipsisme. Je m'oppose à ces arguments, en m'inspirant de travaux récents sur la théorie de la simulation mentale. Je parviens à deux conclusions générales. Premièrement, la méthode de déploiement n'est pas réservée à l'auto-attribution ; elle fonde également l'attribution de croyances à autrui. En second lieu, on peut faire ressortir la spécificité de l'attribution égologique par le « matériau ontologique » auquel cette méthode s'applique. Par exemple, je suis fondé à croire que je crois qu'il pleut parce que c'est le *fait* qu'il pleut, et non une simple possibilité, qui se présente à moi lorsque je me tourne vers le monde. La méthode de déploiement peut échapper au solipsisme si on l'associe à une distinction ontologique naïve entre des faits et de simples possibilités.

# Introspection, déploiement et simulation\*

JÉRÔME DOKIC

EHESS & Institut Jean-Nicod, Paris

dokic@ehess.fr

**RÉSUMÉ.** — Selon une théorie cognitiviste de l'auto-attribution, je peux parvenir à la connaissance directe, non-inférentielle de mes propres croyances. Cette théorie a été traditionnellement associée à la notion d'introspection conçue comme source de connaissance interne. On sait (au moins depuis Wittgenstein) que le recours à cette notion compromet l'application à soi-même d'un concept unifié de croyance, valable également pour autrui. Dans cet essai, j'explore une autre méthode d'auto-attribution, également envisagée par Wittgenstein (et plus tard par Gareth Evans), que j'appelle «méthode de déploiement». Selon cette méthode, je parviens à la connaissance de mes croyances en portant mon attention, non pas à l'intérieur de moi-même, mais directement sur le monde extérieur tel que je l'ai trouvé. Certains arguments wittgensteiniens suggèrent que la méthode de déploiement conduit inexorablement au solipsisme. Je m'oppose à ces arguments, en m'inspirant de travaux récents sur la théorie de la simulation mentale. Je parviens à deux conclusions générales. Premièrement, la méthode de déploiement n'est pas réservée à l'auto-attribution; elle fonde également l'attribution de croyances à autrui. En second lieu, on peut faire ressortir la spécificité de l'attribution égologique par le «matériau ontologique» auquel cette méthode s'applique. Par exemple, je suis fondé à croire que je crois qu'il pleut parce que c'est le *fait* qu'il pleut, et non une simple possibilité, qui se présente à moi lorsque je me tourne vers le monde. La méthode de déploiement peut échapper au solipsisme si on l'associe à une distinction ontologique naïve entre des faits et de simples possibilités.

**ABSTRACT.** — On a cognitivist account of self-ascription, I can have direct, non-inferential knowledge about my own beliefs. This account makes traditionally appeal to the notion of introspection, conceived as an internal source of knowledge. At least since Wittgenstein, many philosophers have justly worried that such a notion makes it impossible to make sense of the ascription of a unified notion of belief, which can be shared with others. In this essay, I explore another method of self-ascription, which was also envisaged by Wittgenstein (and later by Gareth Evans), which I call "the method of deployment". This method is such that I can gain knowledge about my own beliefs by attending not inwards but outwards, on the external world as I found it. Wittgenstein feared that the method of deployment would inevitably lead to solipsism. With reference to recent work in the theory of mental simulation, I try to show that Wittgenstein's fear is ungrounded. I draw two general conclusions. First, the method of deployment does not concern self-ascription only; it can also ground the ascription of beliefs to others. Second, what is special about self-ascription is the "ontological substrate" on which it is based. For instance, I am warranted in believing that I believe that it is raining because I am presented with the *fact* that it is raining, in contrast to a mere possibility. The

---

\* Je remercie Joëlle Proust et Stéphane Chauvier pour leurs commentaires éclairants sur une première version de cet essai.

method of deployment does not lead to solipsism if it is based on a naïve ontological distinction between facts and mere possibilities.

## Introduction<sup>1</sup>

Je crois qu'il pleut, et je sais que je le crois. La connaissance que j'ai de ma croyance semble être directe, non-inférentielle et quasiment infaillible. En tout cas, c'est apparemment d'une toute autre manière que je prends connaissance de la croyance d'autrui. Selon une tradition bien établie, que l'on a coutume de faire remonter à Descartes, je prends connaissance de mes propres états mentaux par introspection, c'est-à-dire par une sorte de sens interne, alors que je parviens à la connaissance des états mentaux d'autrui par observation ou par inférence.

Wittgenstein, on le sait, s'est vigoureusement opposé à la conception cartésienne de l'esprit. Sa critique de l'introspection est cristallisée dans la question rhétorique célèbre, posée dans les *Recherches philosophiques*<sup>2</sup>: « comment puis-je imaginer une douleur que je ne ressens pas sur le modèle d'une douleur que je ressens ? ». De manière générale, la théorie de l'introspection se heurte au problème de l'asymétrie entre la première et la troisième personne, c'est-à-dire entre soi-même et autrui. Il est difficile d'admettre que l'introspection et l'observation, conçues comme des sources de connaissance radicalement différentes, puissent donner accès aux mêmes phénomènes psychologiques. Par conséquent, la méthode introspective compromet l'application à soi-même d'un concept *unifié* de croyance, qui vaut également pour autrui.

Le présent essai ne porte pas directement sur la critique wittgensteinienne de l'introspection<sup>3</sup>, mais sur une autre méthode de connaissance de soi, à première vue plus prometteuse que la méthode introspective. D'un mot, je peux parvenir à la connaissance de mes états mentaux en portant mon attention non pas à l'intérieur de moi-même, mais directement sur le monde extérieur, tel que je le conçois. Si l'on doit la découverte de cette méthode à Wittgenstein (qui s'inspire ici de Moore), le philosophe viennois finira par la tenir pour inadéquate. Les deux méthodes relèvent d'une conception *cognitivist*e de l'auto-attribution, selon laquelle certaines attributions de la forme « Je crois que *p* » sont immédiatement fondées d'un point de vue épistémique et constituent dans le meilleur des cas une *connaissance* de ses propres croyances. Or il est probable que Wittgenstein ait eu davantage de sympathie pour une conception non cognitivist

1. Le présent essai recoupe très largement (avec néanmoins, je l'espère, quelques corrections et précisions) mon article « L'identification de soi, entre savoir-faire et introspection », paru dans les *Cahiers de philosophie de l'Université de Caen* (No 40-41, 2003, p. 45-64), dans un numéro spécial sur la philosophie de Gareth Evans édité par Stéphane Chauvier.

2. Wittgenstein, 1953 : § 302.

3. Pour un exposé magistral de cette critique, cf. Bouveresse (1976). Cf. aussi l'excellent Pouivet (1997).

que *p* » ne sont normalement utilisées ni pour *décrire* ses croyances, ni *a fortiori* pour exprimer une quelconque connaissance à leur sujet<sup>4</sup>.

La méthode anti-introspectionniste rejetée par Wittgenstein a été réactualisée par la philosophie de l'esprit et les sciences cognitives récentes. Il reste à savoir si les objections de Wittgenstein ont été prises en considération. Je tâcherai de montrer que cela n'est pas tout à fait le cas, et que les conceptions cognitivistes actuelles de l'auto-attribution sont au mieux incomplètes.

## 1. Le déploiement contre l'introspection

Dans les *Notes sur l'expérience privée et les « sense data »* (1934-1936), Wittgenstein propose en une brève remarque ce que l'on peut considérer comme une piste originale pour une théorie de l'auto-attribution d'états mentaux :

Si quelqu'un me demande de décrire ce que *je vois*, je décris ce qui *est vu*<sup>5</sup>.

La remarque de Wittgenstein concerne l'expérience visuelle, mais elle s'applique à la croyance, sur laquelle je vais dorénavant me concentrer. Si quelqu'un me demande de décrire ce que *je crois*, je décris ce qui est cru, c'est-à-dire le monde lui-même tel que je le conçois au travers de l'ensemble de mes croyances.

Cette remarque rappelle la doctrine métaphysique du *Tractatus* selon laquelle le sujet n'est pas un objet d'expérience, et ne peut pas être inféré à partir du contenu de l'expérience. « Le monde tel que je l'ai trouvé » est dépourvu de sujet. Le sujet est comme « l'œil géométrique » à l'origine du champ visuel, qui n'est pas lui-même visuellement représenté<sup>6</sup>. Dans les *Notes sur l'expérience privée et les « sense data »*, Wittgenstein fait une application beaucoup plus modeste de cette idée, selon laquelle on peut répondre à une question égologique (« Est-ce que je crois que *p* ? ») à partir d'une expérience dont le contenu (*p*) ne l'est pas forcément.

Comme l'écrit Gareth Evans, qui a contribué de manière significative à l'exploration de la piste ouverte par Wittgenstein, « le point crucial est que [...] lorsque nous nous attribuons une croyance, nos yeux sont pour ainsi dire, et parfois littéralement, dirigés vers l'extérieur – vers le monde<sup>7</sup> ». Evans ne veut pas dire ici que l'on doive chercher à identifier sa croyance au milieu d'autres objets dans le monde. Ce qu'il veut dire, c'est que la réponse que l'on peut donner à une question concernant le monde peut se transformer sans autre forme de procès en une réponse à une question concernant sa croyance :

Si quelqu'un me demande « Pensez-vous qu'il y aura une troisième guerre mondiale ? », je dois porter mon attention, pour répondre à sa question, sur exactement les mêmes phénomènes extérieurs que si je devais répondre à la

4. Sur cette interprétation de Wittgenstein, cf. par exemple Wright (1998).

5. Wittgenstein, 1989, 38.

6. Wittgenstein, 1922, 5, 631-641.

7. Gareth Evans, 1982, 225.

question « Y aura-t-il une troisième guerre mondiale ? ». Je me mets en position de répondre à la question de savoir si je crois que  $p$  en mettant en jeu la procédure, quelle qu'elle soit, dont je dispose pour répondre à la question de savoir si  $p$ <sup>8</sup>.

J'appellerai « méthode de déploiement » la procédure générale, dont Evans décrit ici une application particulière, qui fait dépendre la réponse à une question réflexive de la résolution d'un problème d'ordre inférieur<sup>9</sup>. Cette appellation se justifie par le fait que le sujet déploie le préfixe égologique « je crois que » à partir d'une croyance ou d'une expérience portant sur un état de choses ordinaire, non égologique<sup>10</sup>. Evans considère qu'un examen approfondi de la méthode de déploiement constitue le meilleur remède contre l'introspection, en exorcisant l'idée cartésienne selon laquelle c'est par une sorte de *regard interne* que nous prenons connaissance de nos propres croyances. En effet, l'utilisation de cette méthode ne fait à aucun moment apparaître ma croyance comme un objet donné par une quelconque source de connaissance interne. Les seules sources de connaissance requises pour l'utilisation de la méthode sont celles qui sont de toute façon nécessaires pour fonder la croyance attribuée, à savoir la perception, la mémoire, le témoignage et l'inférence<sup>11</sup>.

Evans fait en outre observer que la méthode de déploiement explique un trait fondamental de beaucoup d'auto-attributions de croyances, à savoir le fait, mis en évidence par Wittgenstein dans le *Cahier bleu*, qu'elles sont protégées contre les erreurs d'identification<sup>12</sup>. Je ne peux pas me tromper sur l'identité de la personne à qui j'attribue une croyance parce que l'utilisation de la méthode de déploiement ne repose pas sur l'identification du sujet de la croyance. Je me contente de suivre la règle suivante, que l'on peut appeler « règle de déploiement » :

---

8. Evans, 1982, 225.

9. Cette appellation correspond à ce que Gordon (1995, 60) nomme *ascent routine*. J'espère que la notion de déploiement ne sera pas confondue avec celle de redéploiement utilisée par Peacocke (1999).

10. Comme le fait observer Peacocke (1999), la procédure d'Evans couvre des cas assez différents. Par exemple, l'auto-attribution d'une croyance que  $p$  peut être fondée sur une attitude occurrente consciente de premier ordre ayant le même contenu (par exemple, un jugement conscient que  $p$ , ou le souvenir apparent que  $p$ ), mais elle peut aussi ne pas être fondée sur une telle attitude. C'est le cas lorsque le sujet doit *se faire un avis* sur la question de savoir si  $p$  (cf. Moran, 1988).

11. Une conséquence plausible de la méthode de déploiement est que l'on ne peut pas identifier « de l'intérieur » une croyance sans en identifier le *contenu*. En général, la méthode ne s'applique qu'aux états mentaux doués de contenu. Je laisse ouverte la question de savoir s'il y a des états mentaux qui n'ont pas de contenu. Cf. Searle (1983, chap. 1) pour une discussion classique de cette question.

12. Wittgenstein (1958] trad. fr. p. 126-127). Sur la notion d'immunité à l'erreur d'identification, cf. aussi les essais réunis dans Shoemaker (1996), à certains desquels Evans se réfère.

[C]haque fois que l'on est en mesure d'affirmer que *p*, on est *ipso facto* en mesure d'affirmer « Je crois que *p* »<sup>13</sup>.

Evans insiste justement sur le fait que la méthode de déploiement ne saurait constituer à elle seule une compréhension adéquate du jugement « Je crois que *p* ». La méthode n'a de sens qu'adossée à un arrière-plan théorique approprié. Parce que la pensée conceptuelle exprimée par « Je crois que *p* », pour une proposition arbitraire *p*, est complexe, elle est soumise à ce qu'Evans appelle la *Contrainte de généralité*. Selon cette contrainte, inspirée de Strawson (1959), les concepts engagés dans une pensée doivent pouvoir figurer indépendamment les uns des autres dans d'autres pensées. Par exemple, il doit être possible de mobiliser le concept de croyance indépendamment du concept égologique, comme dans « Pierre croit que *p* », et le concept égologique indépendamment du concept de croyance, comme dans « Je pèse 72 kilos ».

Selon Evans, qui se démarque ici explicitement de Wittgenstein, la présence d'un arrière-plan théorique approprié rend possible la fonction référentielle du pronom « je », tel qu'il est introduit par la méthode de déploiement. Sans arrière-plan théorique, le « je pense » reste purement formel. Evans considère que « l'ajout de l'arrière-plan ne fait aucune différence pour la méthode d'auto-attribution : en particulier, nous n'avons toujours nul besoin de l'idée d'un regard interne<sup>14</sup> ». La méthode de déploiement nous fait parvenir à la *connaissance* de nos croyances, mais seulement si nous disposons d'un *concept* unifié de croyance, qui s'applique aussi bien à nous-mêmes qu'à autrui. Comme Evans l'écrit à propos du cas où l'on s'attribue une expérience visuelle :

Aucun jugement n'aura le contenu d'une auto-attribution psychologique si le sujet du jugement ne peut pas être considéré comme s'attribuant une propriété qu'il est capable de concevoir comme étant satisfaite par quelqu'un qui n'est pas nécessairement lui-même – un état de choses qu'il devra concevoir comme impliquant un sujet d'expérience persistant dans le temps. Il peut savoir qu'un état de choses de ce type est réalisé simplement en étant conscient d'un arbre, mais il doit concevoir l'état de choses qu'il sait ainsi être réalisé comme étant précisément un état de choses de ce type. Cela signifie qu'il doit se concevoir lui-même, le sujet à qui la propriété est attribuée, comme un être du même genre que lorsqu'il envisage simplement *quelqu'un* en train de voir un arbre – c'est-à-dire, un sujet d'expérience localisé dans l'espace et persistant dans le temps<sup>15</sup>.

Si la méthode de déploiement doit figurer dans une théorie générale de la connaissance de ses propres états mentaux, il faudra résoudre deux problèmes majeurs que je ne puis ici qu'évoquer. Premièrement, il faudra expliquer

13. Evans, 1982, 225.

14. *Ibid.*, 1982, 226.

15. Evans, 1982, 232; c'est moi qui souligne.

les contre-exemples apparents à la méthode de déploiement. Il est difficile, mais pas impossible, d'imaginer qu'un sujet donne une réponse différente à la question de savoir si  $p$  et à la question de savoir s'il croit que  $p$ . Par exemple, il ne perçoit aucune animosité envers son père mais déclare croire qu'il en a à la suite d'une séance avec son psychanalyste<sup>16</sup>. Il est plausible de supposer que de tels contre-exemples (qui, soit dit en passant, sont également des contre-exemples potentiels au paradoxe de Moore: « Je crois que  $p$ , mais  $\neg p$  ») sont nécessairement exceptionnels.

En second lieu, il faudra montrer comment généraliser la méthode de déploiement à d'autres types d'états mentaux. Je supposerai avec Evans que le cas de la croyance est exemplaire, et qu'un modèle adéquat de la connaissance que nous avons de nos propres croyances contient les éléments essentiels pour rendre compte de la connaissance que nous avons de nos désirs, émotions et sensations corporelles, entre autres<sup>17</sup>.

## 2. Le défi solipsiste

Wittgenstein avait manifestement des réserves à l'égard de la méthode de déploiement et était sans doute sceptique quant à sa capacité d'entrer en concurrence sérieuse avec la méthode introspective<sup>18</sup>. Certes, la théorie du déploiement nous dispense de postuler une source de connaissance spéciale, comme l'introspection. Mais elle retient l'idée d'une asymétrie entre la connaissance de nos propres croyances et celle des croyances d'autrui. Wittgenstein identifie l'origine du problème dans ce qu'il appelle l'« énigme solipsiste ». Au sens traditionnel, le solipsisme est la thèse selon laquelle rien n'existe en dehors de moi et de mes états mentaux. La pertinence de cette thèse pour notre propos est qu'elle semble *justifier* de manière particulièrement directe la méthode de déploiement. Considérons une transition particulière de la croyance qu'il pleut à la croyance que je crois qu'il pleut. À première vue, cette transition ne relève pas d'une inférence ordinaire, qui reposerait sur des relations logiques ou probabilistes entre des *contenus* de croyance: le fait qu'il pleut, si c'est un fait, n'implique nullement que je le crois. Comment une telle transition peut-elle alors avoir une quelconque valeur épistémique? Evans cite l'observation suivante attribuée à Wittgenstein par les auteurs du *Wittgenstein Workbook*:

Si quelqu'un me dit, en regardant le ciel, « Je pense qu'il va pleuvoir, donc j'existe », je ne le comprends pas<sup>19</sup>.

16. Pour d'autres considérations de ce genre sur les limites de la méthode de déploiement, cf. Moran (1988), Peacocke (1999) et Martin (1998).

17. Ce qui est en question ici est la croyance pleine, dont la probabilité subjective est 1. Un exposé plus complet devrait prendre en considération l'application de la méthode de déploiement aux croyances partielles. Par exemple, il n'est pas évident que si je crois que  $p$  au degré  $n$ , je doive croire que je crois que  $p$  au même degré.

18. Dans la remarque reproduite au début du §1, les guillemets sont de Wittgenstein, ce qui indique au minimum une prise de distance.

19. Coope, Geach, Potts et White, 1971, 21.

Evans lui-même exprime la même perplexité lorsqu'il écrit :

Comment peut-il se faire que l'on ait connaissance d'un état de choses impliquant un moi substantiel et persistant dans le temps, simplement en étant conscient (pire, en étant seulement apparemment conscient) d'un état du monde<sup>20</sup> ?

Ce que Wittgenstein a perçu nettement, c'est que le solipsisme incarne la tentative la plus directe (et peut-être la seule) de rendre *intelligibles*, d'un point de vue épistémique, les instances de déploiement. Pour le solipsiste, en effet, la transition d'une expérience particulière au jugement que *je suis* le sujet de cette expérience est garantie par le fait que les seules expériences qui existent sont les miennes. Dans le cas de la croyance, la transition d'une croyance que *p* à la croyance que je crois que *p* est justifiée parce que le monde tel que je le conçois, c'est-à-dire tel qu'il est identifié par mes croyances, est nécessairement *mon* monde. Dans la mesure où le monde lui-même, qu'il soit réel ou apparent, porte ma marque, il est possible d'inférer que je crois que *p* à partir du seul *contenu* de ma croyance de premier ordre.

Wittgenstein n'invoque pas ici le solipsisme comme une thèse indépendamment plausible<sup>21</sup>. Par exemple, on peut douter qu'il soit en mesure de reconnaître la distinction entre le monde réel et le monde tel qu'il nous apparaît, et partant la notion d'une croyance vraie ou fausse. Il faut plutôt considérer le solipsisme comme une objection contre la méthode de déploiement, ou plus précisément comme un défi lancé aux cognitivistes qui la croient capable de détrôner la méthode introspective. En effet, si le solipsisme est la seule manière de justifier la méthode de déploiement, on voit mal comment un concept unifié de croyance pourrait en émerger et donc, de manière plus fondamentale, comment une place pourrait être ménagée pour autrui.

Il est facile de sous-estimer le défi solipsiste dans la mesure où il existe apparemment une manière simple d'y répondre. Le fait que je crois que *p* constitue une *raison externe* pour moi de croire que je crois que *p*, puisqu'il garantit objectivement la vérité de la croyance réflexive. Pour un sujet qui sait ce que « Je crois que *p* » veut dire, cette raison externe peut être transformée en *raison interne* : je reconnais que j'ai le droit de croire que je crois que *p* si je crois que *p*. Si le déploiement ne peut pas être conçu sur le modèle d'une inférence ordinaire, à moins de retomber dans le solipsisme, il relève d'un type fiable de transition entre états mentaux, comme le sujet peut le reconnaître sur le plan réflexif.

À mon avis, cette réponse au problème de la justification du déploiement, dont on trouve une forme particulièrement sophistiquée chez Peacocke<sup>22</sup>,

20. Evans, 1982, 231.

21. La question du solipsisme chez Wittgenstein est complexe. Il semble que Wittgenstein reste attaché à une forme de solipsisme dans les *Notes sur l'expérience privée et les "sense-data"*. Il s'agit toutefois davantage d'un solipsisme « méthodologique » que du solipsisme « transcendantal » du *Tractatus*. Sur cette distinction et l'évolution de l'attitude de Wittgenstein à l'égard du solipsisme, cf. Glock (1996) et Pears (1993).

22. Peacocke, 1999, § 5.4.



reste superficielle. Elle suppose à juste titre que la méthode de déploiement est intelligible aux yeux d'un sujet réflexif qui possède déjà le concept égologique. Mais c'est en amont que se situe le problème du solipsisme. Si le concept égologique est au moins partiellement *défini* par la règle de déploiement, comment garantir l'introduction d'un concept unifié de croyance, qui vaut pour soi-même comme pour autrui ?

Le défi solipsiste peut être présenté sous la forme d'un dilemme. D'un côté, la théorie du déploiement, telle qu'on la trouve chez Evans et (sous des formes différentes) dans des travaux plus récents<sup>23</sup>, semble être la seule opposition cognitiviste valable à la théorie de l'introspection (si l'on fait exception du béhaviorisme). D'un autre côté, la seule manière de justifier la méthode de déploiement, en la faisant apparaître comme intelligible aux yeux du sujet, semble impliquer une concession telle au solipsisme que l'introduction d'un concept unifié de croyance risque d'être compromise.

Il est probable qu'un dilemme de ce genre a joué un rôle important dans la réflexion de Wittgenstein sur la nature de l'auto-attribution. L'une des manières de résoudre ce dilemme consiste à rejeter l'un de ses présupposés, à savoir le cognitivisme. La règle de déploiement ne constitue pas une base pour l'introduction d'un concept égologique, car un tel concept n'existe pas. Dans la mesure où « je » n'est pas un terme référentiel, il n'exprime aucun concept ou mode de présentation de soi-même. Par conséquent, « Je crois que *p* » relève moins de la connaissance que de l'*expression* de sa propre croyance. Si l'une des élèves de Wittgenstein, Elizabeth Anscombe, a défendu franchement cette conception<sup>24</sup>, Wittgenstein l'a au moins envisagée, par exemple lorsqu'il suggère dans le *Cahier bleu* qu'aucune référence à soi-même n'est faite dans le langage ou dans la pensée lorsque les mots « Je crois qu'il pleut » sont employés dans une affirmation en lieu et place des mots « Il pleut ».

Dans le reste de cet exposé, j'aimerais indiquer une autre manière possible de résoudre ce dilemme, inspirée par les sciences cognitives récentes et en particulier par la théorie de la simulation mentale. La conception que je vais esquisser retient le présupposé selon lequel le concept égologique est enraciné dans la méthode de déploiement et propose une justification internaliste de cette méthode qui, tout en se tenant à l'écart du solipsisme, lui reconnaît une part de vérité.

### 3. La théorie de la simulation mentale

Selon la théorie de la simulation mentale, la faculté empathique de « se mettre à la place d'autrui », c'est-à-dire d'adopter, sur le mode de l'imagination, la perspective cognitive qu'une autre personne a sur le monde, est essentielle à

23. Outre Gordon et Peacocke, déjà cités, cf. la notion de *perception déplacée* introduite par Dretske (1995 : ch. 2). Cette notion est étroitement liée à la méthode de déploiement, quoique l'analyse de Dretske soit liée à une épistémologie externaliste du déploiement, alors que je cherche à savoir comment le déploiement peut apparaître comme intelligible *aux yeux du sujet*.

24. Cf. Anscombe, 1975.

la compréhension et à l'attribution de ses états mentaux<sup>25</sup>. Dans *Le mot et la chose*, Quine présente de la manière suivante l'intuition qui motive cette théorie :

[D]ans la citation en discours indirect, nous nous projetons nous-mêmes dans ce que nous imaginons qu'a été l'état d'esprit du locuteur, état d'esprit que nous supputons à partir de ses propos et d'autres indices, et ensuite nous disons, dans notre langage, ce qui est naturel ou pertinent dans l'état ainsi feint. [...] Ce qui est en jeu ici, c'est une évaluation, par rapport à certains objectifs particuliers, d'un acte essentiellement dramatique. [...] Nous nous projetons dans l'état d'esprit d'une souris tel que nous l'imaginons à partir de son comportement, et nous le dramatisons en cet état d'esprit sous les traits d'une croyance verbalisée comme il nous paraît pertinent et naturel de le faire dans l'état ainsi feint<sup>26</sup>.

Considérons de plus près la manière dont la capacité d'imaginer des situations contrefactuelles peut être exploitée pour imaginer le monde à partir d'autres perspectives que la sienne propre – non pas le monde tel que je l'ai trouvé, mais le monde tel que d'autres l'ont trouvé, ou tel que je pourrais le trouver. Un enfant joue à téléphoner avec une banane. Il maintient simultanément à l'esprit deux « modèles mentaux » ou systèmes de représentation distincts : son *modèle doxastique*, c'est-à-dire le modèle du monde tel qu'il le croit être, et un *modèle imaginaire*, c'est-à-dire le modèle d'un monde possible distinct du monde doxastique. Dans le premier modèle, l'enfant tient une banane, mais dans le second modèle, il tient un téléphone. Un enfant suffisamment compétent est aussi capable d'*enchâsser* le modèle imaginaire dans le modèle doxastique et de former une représentation plus complexe, que l'on peut rendre de la manière suivante :

Je tiens dans la main une banane, mais dans le jeu, je tiens un téléphone.

Cette représentation, qui fait partie du modèle doxastique du sujet, contient une autre représentation relevant d'un modèle imaginaire *rapporté* à un jeu particulier de faire-semblant (« dans le jeu »). Elle ne repose pas seulement sur la capacité de prévenir toute confusion entre des événements qui se produisent dans des modèles différents ; elle engage surtout la capacité d'établir des comparaisons *à travers les modèles*, et d'apprécier les conséquences que peuvent avoir dans un modèle les événements qui se produisent dans un autre modèle, et vice-versa.

De même qu'une situation imaginaire peut être rapportée à un jeu particulier de faire semblant (« Dans le jeu, *p* »), elle peut être rapportée à une personne particulière (« Selon Pierre, *p* »). On peut ainsi enchâsser une perspective étrangère dans notre perspective propre, sans risque de contradiction.

25. Pour une présentation critique, en français, de la théorie de la simulation, cf. Proust (2001).

26. Quine, 1960-1977, 304-305.

Même si je crois qu'il pleut, je peux imaginer le monde tel qu'il serait s'il faisait beau. Mais je peux aussi rapporter ce monde imaginaire à une personne particulière, par exemple Pierre. J'imagine alors le monde tel que Pierre peut le concevoir. Je forme une représentation complexe, que l'on peut rendre de la manière suivante:

Il pleut, mais selon Pierre, il fait beau.

Si je n'avais pas la capacité de rapporter un monde imaginaire à une personne particulière, je pourrais me représenter par hasard le monde tel que Pierre l'a trouvé sans en apprécier les conséquences dans le monde réel tel que je le conçois. Cette capacité est donc essentielle à la possibilité de la simulation *mentale*<sup>27</sup>. La question pertinente est plutôt de savoir si un enfant qui ne dispose pas encore du concept de croyance peut comprendre une représentation de la forme « Selon Pierre, il pleut », ou si au contraire elle n'est accessible qu'à celui qui maîtrise déjà ce concept. À tout le moins, cette représentation semble décrire le même état de choses que celui qui est décrit par « Pierre *croit* qu'il fait beau ». Indépendamment de cette question, je supposerai que la capacité d'exprimer une conception étrangère du monde sur le mode du « comme si » est un élément central, parmi d'autres, de la possession du concept de croyance.

#### 4. Le déploiement justifié

En ce qui nous concerne ici, la théorie de la simulation mentale présente deux intérêts majeurs. Premièrement, elle permet d'évacuer toute référence à la notion d'introspection (nonobstant Goldman 1992, 1995, 2002). La méthode introspective est effectivement remplacée par la méthode de déploiement, qui fait l'objet d'une utilisation inédite. Nous avons considéré jusqu'à présent que celle-ci était réservée à la première personne et ne servait qu'à s'auto-attribuer une croyance. En fait, étant donné la possibilité d'enchâsser des perspectives les unes dans les autres, la méthode de déploiement peut également servir à attribuer une croyance à autrui. Elle présente ainsi un avantage substantiel par rapport à sa rivale introspective. Comme l'écrit Robert Gordon, l'un des architectes de la théorie contemporaine de la simulation mentale:

À la différence de l'introspection, une méthode de déploiement [*an ascent routine*] pour l'identification des croyances serait aussi bien adaptée pour identifier les croyances d'autrui que pour identifier ses propres croyances<sup>28</sup>.

Supposons que je me mette mentalement à la place de Pierre, c'est-à-dire que je *reproduise* par la simulation certains aspects de son état mental (sans les représenter explicitement). Dans le monde de Pierre, je constate qu'il fait

27. Je considère donc que la théorie de la simulation ne peut pas être « radicale » au sens de Gordon (1995, 1996). Pour une évaluation de ce débat, cf. Dokic et Proust (2002), en particulier les contributions de Joëlle Proust et de Pierre Jacob sur les limites de la simulation.

28. Gordon, 1995, 60.

beau. Je peux alors déployer le concept de croyance dans la portée de ma simulation de Pierre, et faire semblant de former le jugement « Je crois qu'il fait beau ». Le concept *je* utilisé ici fait référence à Pierre, dont je joue provisoirement le rôle. C'est donc à Pierre que j'ai attribué la croyance qu'il fait beau, en utilisant la même méthode que celle qui me permet de m'attribuer la croyance qu'il pleut<sup>29</sup>.

En autorisant l'application de la méthode de déploiement à l'intérieur d'un modèle imaginaire rapporté à une personne particulière, on réduit considérablement le fossé entre la première et la troisième personne et on légitime l'introduction d'un concept unifié de croyance, valable pour soi-même comme pour autrui. Certes, certaines données relatives au comportement et à la situation de Pierre me sont nécessaires pour orienter correctement ma simulation, mais une fois les coordonnées égocentriques effectivement recentrées sur Pierre, mon attribution à celui-ci de la croyance qu'il fait beau a la même forme que mon auto-attribution de la croyance qu'il pleut. En général, la manière dont nous nous représentons les croyances d'autrui et la manière dont nous nous représentons nos propres croyances ont un point essentiel en commun : dans les deux cas, nous portons notre attention non pas sur le monde intérieur du sujet, mais sur le monde extérieur tel qu'il le conçoit.

Le second intérêt majeur de la théorie de la simulation mentale touche plus particulièrement au problème de la justification de la méthode de déploiement. La possibilité même d'exploiter la simulation pour attribuer une croyance à autrui repose sur une dichotomie entre ma perspective doxastique, qui est essentiellement liée aux faits tels que je les conçois, et les autres perspectives, qui peuvent diverger des faits. Cette dichotomie correspond, sur le plan linguistique, à la distinction entre des représentations simples, non enchâssées, et des représentations que l'on ne peut saisir qu'enchâssées dans d'autres représentations. Par exemple, si je crois qu'il pleut, je peux utiliser la représentation « Il pleut » telle quelle dans une affirmation, indépendamment de son enchâssement dans d'autres représentations. Si, par ailleurs, je crois que selon Pierre il fait beau, j'utilise la représentation « Il fait beau » dans la portée de l'opérateur psychologique « selon Pierre ». Certes, je peux utiliser une représentation que je considère comme factuelle dans la portée d'un opérateur psychologique. Par exemple, je peux former une représentation du type « Selon Marie, il pleut ». J'exprime ainsi l'idée que Marie et moi sommes du même avis sur le temps qu'il fait. Nous croyons tous les deux (contre Pierre) qu'il pleut. En règle générale, cependant, l'inférence de « Selon S, *p* » à « *p* » n'est pas valide. On ne peut pas passer de « Selon Marie, il pleut » à « Il pleut » sans faire intervenir quelque autre prémisse.

29. Certes, je ne peux pas exprimer mon attribution à Pierre de la croyance qu'il pleut par les mots « Selon Pierre, je crois qu'il pleut », car le *pronom* « je » est (en français) un terme directement référentiel : il fait référence à moi-même, quels que soient les opérateurs qui, comme « croire », le gouvernent. Je suppose que la thèse de Gordon est plutôt que l'on peut faire *comme si* on utilisait un concept égologique pour désigner une personne qui est en fait quelqu'un d'autre.

La distinction entre ma perspective doxastique et les autres perspectives est étroitement liée à une *ontologie disjonctive*: ce qui correspond dans le monde à une représentation propositionnelle quelconque peut être conçu, ou bien comme un *fait*, ou bien comme une *simple possibilité*. Ce qui est un fait est conçu comme étant essentiellement un fait, alors qu'une simple possibilité est conçue comme étant essentiellement possible. Une telle ontologie est compatible avec l'existence d'une *asymétrie* entre faits et simples possibilités. Dans la théorie de la simulation, une simple possibilité est conçue comme un fait simulé. Il reste qu'un fait simulé est essentiellement différent d'un fait non simulé.

Ma perspective doxastique se distingue de toutes celles que je suis capable d'imaginer par une propriété externe, d'ordre fonctionnel: c'est la seule qui, en général, conduise *effectivement* à l'action et engage mes émotions, douleurs et sentiments *réels*. La suggestion envisagée ici est qu'elle se distingue aussi par une propriété interne, accessible au sujet: c'est la seule qui soit effectivement liée aux faits tels que je les conçois. Certaines conditions subpersonnelles, dont l'étude relève de la neuropsychologie, doivent être remplies pour que les modèles imaginaires ne soient pas confondus avec le modèle de la réalité. Elles ne sont pas remplies dans certains cas pathologiques, comme peut-être la schizophrénie<sup>30</sup>. Or si ces conditions sont inaccessibles au sujet, elles sous-tendent une distinction ontologique consciente, entre des faits et de simples possibilités.

Ces quelques remarques nous permettent d'envisager à nouveaux frais le problème de la justification de la méthode de déploiement. Croire que *p*, c'est agir comme si *p* était vrai, et donc comme si *p* était un *fait*. Lorsque je crois qu'il pleut, la pluie m'apparaît autrement que comme une simple possibilité, ce qui explique le caractère paradoxal de l'énoncé « Je crois qu'il pleut, mais ce n'est pas un fait qu'il pleut ». Or c'est parce que la pluie m'est présentée, sur un plan phénoménologique, comme un fait, que je suis fondé à croire que c'est *moi* qui crois qu'il pleut, plutôt que qui que ce soit d'autre. La logique interne de la simulation garantit l'originalité de la perspective qui est la mienne. Ma perspective est la seule qui n'a pas besoin d'être enchâssée dans d'autres perspectives. C'est la distinction entre les perspectives enchâssées et les autres qui rend possible l'introduction du concept égologique.

Si notre ontologie naïve est disjonctive, au sens où un fait se distingue essentiellement d'une simple possibilité, elle n'est pas solipsiste. Premièrement, comme nous l'avons vu, rien n'empêche qu'autrui partage mes croyances sur la réalité. Ma perspective doxastique n'est pas celle d'un monde privé. Deuxièmement, la transition de ma croyance que *p* à la croyance que je crois que *p* n'est pas une inférence ordinaire, comme dans le cas du solipsisme. C'est seulement une transition fiable entre croyances, mais dont le principe

30. Cf. Currie (2000) et Currie & Ravenscroft (2002, chap. 8).

est directement accessible au sujet. Le monde lui-même n'est pas égologiquement marqué, mais il y a quelque chose dans le *contenu* de mon expérience lorsque je crois que  $p$  qui légitime ma croyance que c'est *moi* qui crois que  $p$ . Lorsque mon expérience me présente le fait que  $p$ , par opposition à une simple possibilité, j'ai une raison interne de croire que je crois que  $p$ . Ma raison n'est pas inférentielle, car je ne suis pas obligé de concevoir le lien entre la factualité de  $p$  et ma croyance que  $p$  sur le modèle d'une implication logique, valant dans tous les mondes possibles. La règle de déploiement que je suis – « Si  $p$  se présente comme un fait, je crois que  $p$  » – ne s'applique pas aux faits considérés comme étant seulement possibles. Ainsi, je suis libre d'imaginer une situation dans laquelle c'est un fait que  $p$ , alors que je ne le crois pas. Par exemple, je peux former la représentation « Il est possible qu'il pleuve mais que je ne le croie pas ». Enfin, pour qui possède un concept authentique de croyance, la transition de « Je crois que  $p$  » à «  $p$  » n'est pas logiquement garantie. Je peux ainsi concevoir la possibilité que ma croyance présente soit fausse. Par exemple, je peux former la représentation « Il est possible qu'il fasse beau mais que je croie qu'il pleut. »

## Conclusion

Si la conclusion à laquelle je suis parvenu n'est pas wittgensteinienne, mon argument prend au sérieux un défi dont Wittgenstein a le mérite d'avoir montré l'importance, à savoir le défi solipsiste. Refusant finalement de faire la moindre concession au solipsisme, Wittgenstein fut enclin à conclure que le pronom « je » (dans son utilisation « subjective ») ne désigne aucun individu, ni sur le plan linguistique ni en pensée. Pour éviter cette conclusion, qui me paraît prématurée, j'ai préféré faire une concession substantielle au solipsisme : ma perspective cognitive sur la réalité est bel et bien essentiellement privilégiée par rapport à d'autres perspectives possibles. Je sais que *je* crois que  $p$  parce que c'est le *fait* que  $p$ , et non pas une simple possibilité, qui m'apparaît lorsque je me tourne vers le monde. Seulement, cette concession peut être détachée de l'affirmation même du solipsisme, selon laquelle rien n'existe en dehors de moi et de mes expériences.

Le modèle épistémologique revendiqué pour les instances de déploiement se situe entre deux extrêmes. D'un côté, le solipsisme opère avec une conception purement internaliste de la justification de telles instances, comme une relation entre des contenus de croyances, en l'occurrence le contenu de la croyance que  $p$  et celui de la croyance que je crois que  $p$ . À l'autre extrême se place une conception purement externaliste de la justification des instances de déploiement, selon laquelle la transition entre la croyance que  $p$  et la croyance réflexive correspondante est justifiée seulement en tant qu'elle relève d'une transition fiable entre états mentaux. La position intermédiaire que j'essaie de promouvoir est externaliste au sens où la règle de déploiement garantit objectivement la vérité de la croyance réflexive étant donné la croyance de première ordre, mais elle incorpore un élément internaliste, car

le sujet, loin de suivre la règle aveuglément, fonde son application sur la manière dont *p* apparaît au sujet, c'est-à-dire (quand celui-ci croit que *p*) comme un fait.

Si la voie esquissée dans cet essai est possible, nous sommes en droit d'espérer une généalogie du concept égologique et de la signification du pronom « je » qui ne repose ni sur la présence du sujet comme un élément de l'expérience ni sur toute autre représentation égologique non conceptuelle<sup>31</sup>, mais sur une distinction ontologique primitive entre faits et simples possibilités. Cette distinction fournit une base indépendante pour l'introduction du concept égologique et du pronom « je ». Elle constitue ainsi l'un des fondements de la subjectivité<sup>32</sup>.

## Références

- Anscombe, G. E. M. « The First Person », in S. Guttenplan (dir.), *Mind and Language*, Oxford, Oxford University Press, 1975.
- Bermúdez, J. *The Paradoxes of Self-Consciousness*, Cambridge (Mass.), MIT Press, 1998.
- Bouveresse, J. *Le mythe de l'intériorité. Expérience, signification et langage privé chez Wittgenstein*, Paris, Minuit, 1976.
- Carruthers, P. et P. K. Smith (dir.), *Theories of theories of mind*, Cambridge, Cambridge University Press, 1996.
- Chauvier, S. *Dire « Je »*, Paris, Vrin, 2001.
- Coope, C., P. Geach, T. Potts et R. White (dir.), *A Wittgenstein Workbook*, Oxford, Blackwell, 1971.
- Currie, G. « Imagination, Delusion and Hallucinations », in M. Coltheart et M. Davies (dir.), *Pathologies of Belief*, Oxford, Blackwell, 2000.
- Currie, G. et I. Ravenscroft. *Recreative Minds*, Oxford, Clarendon Press, 2002.
- Davies, M. et T. Stone (dir.), *Folk Psychology*, Oxford, Blackwell, 1995a.
- . *Mental Simulation*, Oxford, Blackwell, 1995b.
- Dokic, J. et J. Proust (dir.), *Simulation and Knowledge of Action*, Amsterdam/Philadelphia, John Benjamins, 2002.
- Dretske, F. *Naturalizing the Mind*, Cambridge (Mass.), MIT Press, 1995.
- . *Perception, Knowledge and Belief*, Cambridge, Cambridge University Press, 2000.
- Engel, P. *Philosophie et psychologie*, Paris, Folio-Gallimard, 1996.
- Evans, G. *The Varieties of Reference*, Oxford, Clarendon Press, 1982.
- Glock, H.-J. « Solipsism », in *A Wittgenstein Dictionary*, Oxford, Blackwell, 1996.
- Goldman, A. « In defense of the simulation theory », *Mind and Language*, 7, 1992, p. 104-119.
- . « Interpretation Psychologized », in Davies et Stone (dir.), 1995a.

31. Mon argument n'exclut cependant pas qu'il y en ait une. Cf. Bermúdez (1998), qui défend l'existence de représentations égologiques non conceptuelles.

32. Une conséquence de ma thèse est que l'utilisation intelligente du pronom « je » ne saurait constituer la subjectivité, comme le propose Chauvier (2001). Mes quelques remarques ne rendent pas justice à son remarquable essai, même si, pour autant que je puisse en juger, il néglige la position esquissée ici.

- . « Simulation and Mental Concepts », in Dokic et Proust (dir.), 2002.
- Gordon, R. 1995. « Simulation Without Introspection or Inference from Me to You », in Davies et Stone (dir.), 1995b.
- . « “Radical” simulationism », in Carruthers et Smith (dir.), 1996.
- Heal, J. 1995. « How to Think About Thinking », in Davies et Stone (dir.), 1995b.
- . « Simulation, theory, and content », in Carruthers et Smith (dir.), 1996.
- Jacob, P. « The scope and limits of mental simulation », in Dokic et Proust (dir.), 2002.
- Lewis, D. *On the Plurality of Worlds*, Oxford, Blackwell, 1986.
- Martin, M. G. F. « An Eye Directed Outward », in Wright, Smith et Macdonald (dir.), 1998.
- Moran, R. « Making Up Your Mind: Self-Interpretation and Self-Constitution », *Ratio* I, 2, 1988, p. 135-151.
- Peacocke, C. *Being Known*, Oxford, Clarendon Press, 1999.
- Pears, D. *La pensée-Wittgenstein: du Tractatus aux Recherches philosophiques*, 1993; trad. fr. C. Chauviré, Paris, Aubier, 1993.
- Pollock, J.-L. et J. Cruz. *Contemporary Theories Of Knowledge*, Rowman & Littlefield Publishers, 1999.
- Pouivet, R. *Après Wittgenstein, saint Thomas*, Paris, PUF (collection « Philosophie »), 1997.
- Proust, J. « Pour une théorie “motrice” de la simulation », *Psychologie française* 45, 4, Numéro spécial sur la simulation de l'action (éd. par J. Decety), 2001, p. 295-306.
- . « Can “Radical” Simulation Theories Explain Psychological Concept Acquisition? », in Dokic et Proust (dir.), 2002.
- Quine, W. V. O. *Word and Object*, Cambridge (Mass.), MIT Press, 1960; trad. fr. J. Dopp et P. Gochet, *Le mot et la chose*, Paris, Flammarion, 1977.
- Searle, J. *Intentionality*, Cambridge, CUP, 1983; trad. franç. C. Pichevin, *L'intentionnalité*, Paris, Minuit, 1987.
- Shoemaker, S. *The First-Person Perspective and Other Essays*, Cambridge, CUP, 1996.
- Strawson, P. F. *Individuals*, Londres, Methuen, 1959; trad. franç. A. Shalom et P. Drong, *Les Individus*, Paris, Seuil, 1973.
- Wittgenstein, L. *Tractatus Logico-Philosophicus*, London, Routledge & Kegan Paul, 1922; trad. fr. G.-G. Granger, Paris, Gallimard, 1993.
- . *Philosophical Investigations*, éd. par E. Anscombe et R. Rhees, Oxford, 1953; trad. franç. P. Klossowski, Paris, Gallimard, 1961.
- . *The Blue and Brown Books*, éd. par R. Rhees, Oxford, 1958; trad. fr. G. Durand, *Le Cahier bleu et le Cahier brun*, Paris, Gallimard, 1965.
- . *Notes sur l'expérience privée et les « sense data »* (1934-1936), 1989; trad. franç. E. Rigal, 2<sup>e</sup> édition, Mauvezin, TER.
- Wright, C. « Self-Knowledge: The Wittgensteinian Legacy », in Wright, Smith et Macdonald (dir.), 1998.
- Wright, C., B. C. Smith, et C. Macdonald, (dir.), *Knowing Our Own Minds*, Oxford, Clarendon Press, 1998.